

BAKER BOTTS LLP

910 LOUISIANA
HOUSTON, TEXAS
77002-4995

TEL +1 713.229.1234
FAX +1 713.229.1522
BakerBotts.com

AUSTIN
BRUSSELS
DALLAS
DUBAI

HOUSTON
LONDON

NEW YORK
PALO ALTO
RIYADH
SAN FRANCISCO
SINGAPORE
WASHINGTON

April 7, 2023

Hon. Andrew M. Edison
601 Rosenberg, Seventh Floor
Galveston, Texas 77550

Amy Pharr Hefley
TEL +1 713.229.1270
FAX +1 713.229.2770
amy.hefley@bakerbotts.com

Dear Judge Edison,

Defendants have agreed to produce documents in response to 55 requests for production, spanning a six- to eight-year period, relating to the Alpine High Project run out of Apache's San Antonio office, which closed down approximately three years ago.¹ Defendants have agreed to 60 document custodians (many added at Plaintiffs' request) and to separately search non-custodial central files and document repositories. Defendants have proposed search terms they know will capture hundreds of thousands of documents and gave Plaintiffs an opportunity to comment on them, and Defendants stand ready to begin reviewing and producing the documents that hit on those terms. Defendants are confident their plan is a reasonable way to capture potentially relevant documents. But they also have made clear that they will remain open to refining their searches as discovery progresses, including in response to any concerns Plaintiffs may raise after receiving and reviewing Defendants' rolling productions.

What has bogged down that process is the approach exemplified by Plaintiffs' letter (Dkt. 100), which seeks to (1) second-guess and overhaul Defendants' methodology for collecting and reviewing their own electronically stored information ("ESI"), and (2) wage hypothetical, preemptive battles over production disputes that do not—and will likely never—exist. *See Nichols v. Noom Inc.*, 2021 WL 948646, at *5 (S.D.N.Y. Mar. 11, 2021) (noting "inherent tension" between "address[ing] all potentially foreseeable [ESI] issues ... and getting discovery underway") (attached as Ex. 1). As this Court has held, the Sedona Principles counsel firmly against such efforts to "micro-manage the parties' internal review procedures." *Edwards v. McDermott Int'l, Inc.*, 2021 WL 5121853, at *3 (S.D. Tex. Nov. 4, 2021). The Court should reject Plaintiffs' requests.

I. Plaintiffs' requested ESI Order is neither necessary nor proper.

After multiple meet-and-confers on an agreed ESI Protocol—and multiple rounds of revisions and concessions by both sides—the "near-standstill" Plaintiffs describe (Ltr. 1) is due to Plaintiffs' unwillingness to budge on a handful of unreasonable demands. Plaintiffs now ask this Court not only to foist those terms on Defendants, but also to sign the *entirety* of Plaintiffs' ESI Protocol as an "ESI Order" of the Court over Defendants'

¹ Although this is, at its core, a discovery dispute manufactured by Plaintiffs, Defendants also take issue with Plaintiffs' unnecessary and inaccurate characterization of the case. The Court already described this case's surviving dismissal as a "very close call," and Defendants look forward to refuting Plaintiffs' allegations.

BAKER BOTTS LLP

Hon. Andrew M. Edison

- 2 -

April 7, 2023

objection. *Id.* at 3. That request is inappropriate on both a macro and micro level.

For starters, an agreed ESI Protocol is just that—an *agreement* between the Parties on a framework to guide the ESI discovery process, consistent with this Court’s stated preference for parties “to work together in a cooperative fashion” on ESI issues, so that “district court judges” remain able to avoid “micro-manag[ing] the parties’ internal review procedures.” *Edwards*, 2021 WL 5121853, at *3. Defendants told Plaintiffs from the start that Defendants were willing to negotiate an ESI Protocol, but only as an agreement between the Parties, not an order to be presented to the Court, given the Sedona Principles’ clear mandate against such efforts to micro-manage the process. Defendants negotiated on that basis, only to have Plaintiffs now ask the Court to enter an (objectionable) ESI Order. Plaintiffs identify no case in this District to unilaterally impose one party’s preferred protocol as an ESI Order governing the other, and cases, including complex § 10(b) cases, regularly proceed *without* such orders. As in those cases, Defendants ask the Court to simply address—and, as explained below, reject—each of Plaintiffs’ “unresolved issues,” Dkt. 86, without taking the unnecessary additional step of entering an ESI Order dictating all manner of internal review protocols.

Hyperlinked Documents. Plaintiffs’ request for an order remedying a supposed “refus[al] to collect” hyperlinked email “attachments” is wrong in several ways. Defendants have no qualms (a) “checking the box” in Purview and collecting the linked internal document (though they cannot guarantee the collected document will be “the version as-shared” in the original email, *contra* Ltr. 3), and (b) searching in parallel for the same underlying documents in the shared folders where they might live. But Defendants object to Plaintiffs’ proposed ESI Protocol language because it treats links to non-static documents located elsewhere as “attachments,” when *courts* “do not agree that a hyperlinked document is an attachment.” *Noom*, 2021 WL 948646, at *4 (rejecting preemptive effort to force party to treat hyperlinked documents as attachments in the “email ‘family,’” and allowing such documents to be “produced separately” instead).

Plaintiffs’ cited cases (at Ltr. 3 & n.2) all involve the very different scenario where a party either failed to produce an actual attachment (as opposed to a hyperlinked document) together with its parent email, or resisted a post-production request to search for and produce the underlying document that had been hyperlinked to a produced email despite that request being directed at only a (relatively small) subset of emails. Here, as in *Noom*, “[a]t this point it is entirely speculative how many underlying hyperlinked documents are relevant and material to this case.” 2021 WL 948646, at *4. The issue can and should be addressed as Defendants make their productions, at which time Defendants will entertain requests to search for additional hyperlinked documents that Plaintiffs identify as potentially being responsive. Should “Plaintiffs determine” at a later point in the production process that “there is a need for an additional targeted pull or production or clarifying information about a hyperlinked document’s identity or Bates number,

BAKER BOTTS_{LLP}

Hon. Andrew M. Edison

- 3 -

April 7, 2023

Plaintiffs can request it,” and Defendants will address such requests in due course. *Id.*

Future TAR Optionality. Plaintiffs next propose a categorical rule that a party must “cull responsive documents utilizing either search terms or TAR, but not both.” Ltr. 3. Their only cited case eschews any such general bans, recognizing that cases *do* exist “where parties are permitted to apply TAR after culling by the application of search terms.” *In re Allergan Biocell Textured Breast Implant Prods. Liab. Litig.*, 2022 WL 16630821, at *2 (D.N.J. Oct. 25, 2022).² That is, the use of search terms and TAR is a case-specific matter for which “courts [must] find solution to the problems confronting them.” *Id.*

No such problems currently confront the Court here, because what Defendants propose is not an ESI Protocol that intractably binds Plaintiffs to “whatever search method [Defendants] want” and under whatever circumstances. *Contra* Ltr. 3. Instead, Defendants merely seek to preserve the *option*—based on (a) the volume of documents elicited by the final search-term set, and (b) the status of Defendants’ review efforts—to elect TAR in the future and to be transparent about what that election and resulting TAR Protocol would entail. At *that* point, the Parties would meet and confer, and Plaintiffs could raise any unresolved issues regarding the proposed TAR protocol with the Court. Ex. 2, Defs.’ Redline ESI Protocol.³ Nothing supports Plaintiffs’ effort to preemptively and permanently foreclose that option at this early juncture, before the Parties can even assess the hypothetical cost-benefit analysis that might apply as discovery evolves.

Validation Protocol for Human Review. Defendants agreed to disclose and discuss their quality control measures if they elect to add TAR (machine-based review), so Plaintiffs complain only about the lack of a similarly formal “validation protocol” for search-term based human review. Defendants *agree* that “quality assurance techniques” are important for human review, *In re Seroquel Prods. Liab. Litig.*, 244 F.R.D. 650, 662 (M.D. Fla. 2007), and they of course intend to apply—not “reject”—such QC measures through the multi-level review and sampling processes that are ubiquitous in modern discovery. Plaintiffs go astray by insisting that (a) Defendants’ QC process for human review mimic the detailed validation protocols necessary for TAR, and (b) Plaintiffs be able to probe the minutiae of how outside counsel will structure their multi-level, attorney-review teams. No authority supports either contention. *See Edwards*, 2021 WL 5121853, at *3 (no “discovery on discovery” absent “specific, tangible, evidence-based indicia (versus general allegations of deficiencies or mere ‘speculation’) of a material failure by the

² E.g., *In re Valsartan, Losartan, & Irbesartan Prods. Liab. Litig.*, 337 F.R.D. 610, 616 (D.N.J. 2020) (“Ample case law exists to support Teva’s position that in appropriate instances layering”—“layer[ing] a document production with search terms and TAR”—“may be done.”); *In re Biomet M2a Magnum Hip Implant Prods. Liab. Litig.*, 2013 WL 1729682, at *1-2 (N.D. Ind. Apr. 18, 2013) (permitting TAR after “keyword culling”); *Bridgestone Ams., Inc. v. IBM Corp.*, 2014 WL 4923014, at *1 (M.D. Tenn. July 22, 2014) (“permit[ting] Plaintiff to use predictive coding on the documents that they have presently identified, based on the search terms Defendant provided.”).

³ Exhibit 2 is a version of the ESI protocol from late in the Parties’ negotiations. It contains dozens of concessions made by Defendants in hopes of reaching an agreement. If the Court were inclined to enter an order adopting one side’s preferred ESI protocol, Defendants request an opportunity to submit their preferred version for consideration.

BAKER BOTTS LLP

Hon. Andrew M. Edison

- 4 -

April 7, 2023

responding party to meet its obligations”) (quoting SEDONA PRINCIPLES at 123).⁴

II. Plaintiffs cannot micro-manage Defendants’ proposed search methodology.

Despite touting the technical capabilities of Microsoft’s Purview e-discovery tool a few pages earlier, Plaintiffs’ letter goes on to disparage it as a supposedly non-viable mechanism for collecting documents. Those contentions misapprehend the governing e-discovery principles and severely downplay Defendants’ planned search methodology.

As this Court has recognized, “[r]esponding parties are best situated to evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own [ESI],” and it is not the role of “district court judges”—much less parties—to “micro-manage [other] parties’ internal review procedures.” *Edwards*, 2021 WL 5121853, at *3 (quoting SEDONA PRINCIPLES at 118). While Plaintiffs clearly prefer that Apache apply “date filters to collect documents and then apply[] search terms,” Ltr. 5, Plaintiffs overlook that (1) Plaintiffs’ lengthy search period would require Defendants to bear the substantial cost of processing *six to eight years’ worth of data across 60 custodians*, regardless of its potential relevance; and (2) Apache has routinely and successfully employed Purview on *its systems* for *its document-collection process*—in cases large and small—without experiencing the parade of horrors Plaintiffs posit. Despite a single special master’s criticism of Purview, Ltr. 5, courts in this District are not bound by processes “condoned by other courts, or representative of the [purportedly] highest standards of practice”; rather, “the Court’s decision must ultimately reflect the merits of the situation at hand.” *Dizdar v. State Farm Lloyds*, 2015 WL 12781020, at *11 (S.D. Tex. Jan. 7, 2015). Because “complying with Plaintiffs’ [proposed] ESI Protocol would require [Apache] to re-engineer its processes for retrieving, storing, and producing information for litigation,” that is a compelling reason to reject Plaintiffs’ request. *Id.*

Plaintiffs’ complaint that Purview “does not index all data (including audio and video files),” Ltr. 5, also rings hollow given Defendants’ stated intention to search files (*e.g.*, shared files, the server for the San Antonio office associated with Alpine High, etc.) in addition to the custodial search-term process. Plaintiffs protest that “Defendants have

⁴ Plaintiffs’ effort (at Ltr. 4 n.5) to force Defendants to produce FilePath metadata flunks the proportionality test. Defendants have agreed to produce 41 categories of metadata, which will give Plaintiffs extensive information about every produced document, but object to Plaintiffs’ demand for FilePath metadata on burden and relevance grounds. Producing it would entail the additional time-consuming and costly step of reviewing the FilePath description of *every* produced document for privilege and work-product, despite the comparatively minimal, if any, relevance of such metadata in a case about whether Defendants’ public statements were misleading. That Plaintiffs had to resort to citing a *trade-secrets* case—where “the location and manner in which [the party] stored its ESI [wa]s directly relevant” to whether it “took reasonable measures to maintain the secrecy of its alleged trade secrets,” *Javo Beverage Co., Inc. v. Cal. Extraction Ventures, Inc.*, 2020 WL 2062146, at *8 (S.D. Cal. Apr. 29, 2020)—proves Defendants’ point that FilePath is not a standard metadata field produced in § 10(b) cases. The appropriate way to balance the parties’ competing interests here is what Defendants proposed—Plaintiffs can identify specific, produced documents for which they believe FilePath is relevant, and Defendants can address those specific requests.

BAKER BOTTS LLP

Hon. Andrew M. Edison

- 5 -

April 7, 2023

not disclosed which additional sources they will search,” *id.* at 4-5, but that is because Defendants are still identifying the universe of available non-custodial sources, given that the San Antonio office closed down three years ago (and nearly a year before this suit was filed) and its employees were mostly let go. Plaintiffs cannot weaponize Defendants’ careful efforts to be comprehensive and accurate in their transparency about those non-custodial sources to try to overhaul Defendants’ primary means of custodial document collection. Plaintiffs have not demonstrated that Defendants’ methodological “choice is manifestly unreasonable,” so “the court should play no role in dictating the design of the search,” including by “choosing search tools.” *Mortgage Resolution Servicing, LLC v. JPMorgan Chase Bank, N.A.*, 2017 WL 2305398, at * 2 (S.D.N.Y. May 18, 2017).

Plaintiffs’ complaints about the hypothetical shortcomings of Purview (plus non-custodial source searches), gets the burden backwards. The “*requesting party* has the burden ... to show that the responding party’s steps to preserve and produce relevant [ESI] were inadequate.” SEDONA PRINCIPLES at 131 (emphasis added). The proper path forward is to let Defendants proceed and “if a dispute arises about materials Plaintiffs identify as missing from [the] productions, the parties should mediate it and, if they are not able to resolve it, file an appropriate motion.” *In re Facebook, Inc. Consumer Privacy User Profile Litig.*, 2021 WL 10282172, at *20 (N.D. Cal. Oct. 11, 2021). The solution is not to jettison the collection tool Apache has used in many cases with no issue.

III. Plaintiffs’ remaining, conclusory requests are premature and improper.

Search Terms. Defendants sent revised search terms on March 31, 2023, triggering Plaintiffs’ right to decide if “disputes still exist” and “submit them to the Court.” Dkt. 100-3 ¶ IV.D. Despite conceding they are still reviewing the terms, Plaintiffs gesture vaguely (and misleadingly⁵) at perceived deficiencies. Plaintiffs are welcome to bring actual disputes to the Court by April 28, 2023, but Court-ordered bi-weekly meet-and-confers and status reports would needlessly burden the Parties and the Court.

Fact Discovery Deadline. In response to Plaintiffs’ concerns with the December 1, 2023 fact-discovery deadline (that *Plaintiffs* proposed, Dkt. 85-1), Defendants offered “to discuss with you a modification of that deadline”—an offer Plaintiffs ignored. The Court should instruct the Parties to confer and try to agree on a modification to the schedule.

Rolling Production. Plaintiffs ask that Defendants be ordered to produce documents that Plaintiffs baselessly claim can be found “without resort to ESI methodologies.” Ltr. 1, 5. Defendants already are preparing their next rolling production with a target of next week. There is no basis for a Court-ordered interim production deadline—and certainly not for the ill-defined category Plaintiffs seek.

⁵ For example, Plaintiffs omit that nearly 30 of Defendants’ terms hit on >10,000 documents, with one term hitting on >100,000 documents. Meanwhile, Plaintiffs’ proposed search terms included strings that were extraordinarily overbroad, such as *all emails* between one (non-defendant) custodian and five others over a multi-year period.

BAKER BOTTS LLP

Hon. Andrew M. Edison

- 6 -

April 7, 2023

/s/ Amy Pharr Hefley

BAKER BOTTS L.L.P.

David D. Sterling

Attorney-In-Charge

State Bar No. 19170000

Federal I.D. No. 07079

Amy Pharr Hefley

State Bar No. 24046046

Anthony J. Lucisano

State Bar No. 24102118

Federal I.D. No. 3369146

C. Frank Mace

State Bar No. 24110609

Federal I.D. No. 3385915

910 Louisiana Street

Houston, TX 77002

Telephone: (713) 229-1946

Facsimile: (713) 229-7946

david.sterling@bakerbotts.com

amy.hefley@bakerbotts.com

anthony.lucisano@bakerbotts.com

frank.mace@bakerbotts.com

*Counsel for Defendants Apache
Corporation, John J. Christmann IV,
Timothy J. Sullivan, and Stephen J. Riney*

cc: All Counsel of Record via ECF

EXHIBIT 1

2021 WL 948646

Only the Westlaw citation is currently available.

United States District Court, S.D. New York.

Mojo NICHOLS, et al., Plaintiffs,

v.

NOOM INC., et al., Defendants.

20-CV-3677 (LGS) (KHP)

|

Signed 03/11/2021

Attorneys and Law Firms

J. Burkett McInturff, Jessica Hunter, Steven Lance Wittels, Wittels McInturff Palikovic, Steven Dana Cohen, Wittels Law, Armonk, NY, for Plaintiffs Maryanne Deracleo, Jennifer Sellers, Rebecca Richards, Susan Brewster, Mojo Nichols, Karen Kelly.

J. Burkett McInturff, Jessica Hunter, Steven Lance Wittels, Wittels McInturff Palikovic, Armonk, NY, for Plaintiffs Duane Dea, Stacy Spencer.

Ian Ross Shapiro, Charles Low, Cooley LLP, New York, NY, Aarti Reddy, Colin Sarver Scott, Joseph Mornin, Max A. Bernstein, Michael Graham Rhodes, Cooley LLP, San Francisco, CA, for Defendant Noom Inc.

Aarti Reddy, Colin Sarver Scott, Joseph Mornin, Cooley LLP, San Francisco, CA, Charles Low, Cooley LLP, New York, NY, for Defendant Artem Petakov.

DISCOVERY ORDER CONCERNING PRODUCTION OF HYPERLINKED INTERNAL DOCUMENTS

KATHARINE H. PARKER, United States Magistrate Judge

*1 Plaintiffs filed a letter motion on February 19, 2021 seeking to “clarify” the Court’s previous orders and rulings regarding Defendants’ production of documents linked to Google Drive documents and Gmail communications via hyperlink. (ECF No. 214.) In the alternative, they seek reconsideration of this Court’s prior rulings that (1) Noom would be permitted to utilize Google Vault to collect its Google Drive and Gmail documents and (2) to the extent Plaintiffs identified certain internal documents containing hyperlinks to internal Noom documents that appeared to be material to the claims or defenses in this action and

could not locate the corresponding hyperlinked document in the production, they should raise the issue with Noom and Noom would be required to provide the document or its Bates number (unless withheld on privilege or work product grounds and logged). In this Court’s experience, only a fraction of the documents produced in discovery will be material to the litigation and the Court indicated to the parties that the above-described procedure should be sufficient to address Plaintiffs’ concerns about identifying key hyperlinked documents.

Plaintiffs contend that, based on Noom’s early production of documents, they have learned Noom employees frequently link to internal documents in lieu of attachments to emails or other documents. They argue that hyperlinks are akin to attachments and should be produced as part of a document “family.” They argue that without metadata linking the underlying hyperlinked Noom document to the document containing the hyperlink, they will not be able to determine families of documents. They also express concern that some of the hyperlinked documents may not be produced at all. They urge this Court to require Noom to use MetaSpike’s Forensic Evidence Collector (“FEC”) to recollect Google Drive and Gmail documents so that any hyperlinked documents are also pulled as part of the document “family” or to create a program using Google’s application programming interface to extract links from responsive Google Drive documents, retrieve those linked documents, and produce them as attachments. Plaintiffs estimate it will take only one to two weeks to write a program to extract the links. They do not state what the time or costs would be for processing, de-duplication, and re-review of documents.

Noom opposes Plaintiffs’ motion, arguing that the hyperlinks are not attachments and that it is separately collecting and producing relevant internal documents on Google Drive such that there should be no concern that Plaintiffs will not receive relevant internal documents. Noom has agreed to produce a reasonable number of linked documents at Plaintiffs’ request and has been ordered by this Court at a discovery conference to provide such linked documents to the extent Plaintiffs cannot locate them in the production and/or need clarification as to whether a certain document is the corresponding linked document. Noom’s discovery expert submitted a declaration stating that the FEC tool is unworkable and that Noom would incur roughly \$180,000 in costs to collect the hyperlinked documents and produce them, resulting in further delays. For that reason, it argues that Plaintiffs’ request is not proportional to the needs of the case.

*2 The issues raised by Plaintiffs raise complex questions about what constitutes reasonable search and collection methods in 2021—when older forms of communicating via emails and documents with attachments and footnotes or endnotes are replaced by emails and documents containing hyperlinks to other documents, video, audio, or picture files. It also highlights the changing nature of how documents are stored and should be collected. The Court has carefully considered the issues raised and, for the reasons set forth below, denies Plaintiffs’ motion for reconsideration and trusts that this Opinion provides clarification.

BACKGROUND

This case was filed in May 2020. Both sides have sophisticated counsel knowledgeable about discovery of electronically stored information (“ESI”) and are utilizing experienced ESI consultants to assist with the collection and production of documents. In accordance with best practices, the parties began to negotiate an ESI protocol but disagreed on major aspects of it, requiring multiple conferences before the Court. Ultimately, Noom agreed to collect and did collect relevant data from multiple sources including Gmail, G-chat, Google Drive, Google Calendar, Slack, and several other reporting tools/databases. Noom also agreed to produce a long list of metadata. (ECF No. 146, Section C(6).) On November 13, 2020, the parties advised the Court that they agreed to Noom’s use of Google Vault to collect documents from Google Drive even though “filepath” metadata field would not be available for documents collected that way.¹ (ECF No. 98.)

¹ Filepath information reveals the original location of an email or shows where a document was originally stored. Noom agreed to provide filepath metadata to the extent it could be reasonably collected.

The parties disagreed, however, on the use of Google Vault to collect Gmail emails. Plaintiffs asked Noom to use FEC to collect email, which it contended was a superior tool to Google Vault and no more expensive or onerous. Both sides submitted expert declarations from their ESI consultants about the pros and cons of FEC versus Google Vault. One main concern of Plaintiffs was that a Google Vault collection would not pull documents referenced in emails by a hyperlink.

Noom argued that it should be permitted to determine the method of collection so long as it is reasonable and that, because it was producing actual attachments to emails and separately collecting Google Drive documents that are referenced as file hyperlinks through its custodial and shared drive collection, it should not be required to perform an essentially redundant collection that would be burdensome. That is, Plaintiffs would receive relevant and responsive hyperlinked Google Drive documents from Noom’s separate collection of those documents. One of Noom’s ESI consultants also pointed out that Google Vault is the standard industry tool that is used for Gmail export by the vast majority of its Google business clients. (ECF No. 105-1, Declaration of Jason Sims ¶ 7.) Further, in response to Plaintiffs’ request that it consider using FEC, Noom’s vendor tested FEC and found that it was not workable insofar as the tool “continues to stall, leading to significantly longer processing times that will likely further delay collection.” (ECF No. 105.) Its expert also stated that collection of emails using FEC while also collecting documents in Google Drive would “substantially increase its costs and result in a significant number of duplicates in its review population that cannot be culled through deduplication.” *Id.*

After fully hearing the parties’ arguments, the Court held that Noom could use its preferred software to collect email documents, finding that method reasonable and deferring to the principle that a producing party is best situated to determine its own search and collection methods so long as they are reasonable.² See Sedona Principle 6; *Hyles v. New York City*, 10 Civ. 3119 (AT) (AJP), [2016 WL 4077114](#) (S.D.N.Y. Aug. 1, 2016). The Court also took into account the relative costs and delays attendant to utilizing FEC. However, as noted above, to address Plaintiffs’ concern about not being able to identify which Google Drive documents in the production related to a particular hyperlink, the Court directed that if there were particular key documents containing hyperlinks where the hyperlinked documents could not be located in the production, Plaintiffs could raise that issue with the Court. In a later conference the Court held that if there were certain documents discovered in the production containing hyperlinks for which the corresponding hyperlinked document could not be located or identified, Plaintiffs could raise the issue with Noom and Noom would be required to provide the document or Bates number. Noom was, from the start, willing to do this for a reasonable number of documents.

2 Noom did use FEC to collect Google Calendar ESI.

*3 Notably, the ESI protocol negotiated by the parties and later entered by the Court does not state that hyperlinked documents are part of “family groups.” But the protocol does not define “attachments” either. Section C(5) of that Order, entitled “Family Groups,” says:

(a) A document and all other documents in its attachment range, emails with attachments, files with extracted embedded OLE documents, and email or other documents together with any documents referenced by document stubs within those emails or other documents all constitute family groups. If any member of a family group is produced, all members of that group must be also be produced or else logged as privileged.

(b) Parent-child relationships (the association between an attachment and its parent document or between embedded documents and their parent) shall be preserved.

(ECF No. 146.) It is clear to this Court that there was no meeting of the minds on whether hyperlinks were attachments and this Court, when entering the order, did not view hyperlinks to be attachments. While the protocol does reference “files with extracted embedded OLE documents,” the Court understands this to refer to *embedded*, displayed documents such as a graph or a chart within a Word document or email—not hyperlinked documents. The ESI also refers to “document stubs,” which the Court understands to refer to the company's archive location of attachments to emails.³ In sum, the ESI order does not treat hyperlinked documents as attachments. As noted above, Noom is producing all attachments to emails with the parent email as part of the email “family.” Relevant hyperlinked documents are being produced separately.

3 To the extent Plaintiffs now argue that “stub” was intended to require pulls of hyperlinked documents, this is not the intent Noom or the Court ascribed to that term.

DISCUSSION

The Court may grant a motion for reconsideration when there has been an intervening change of controlling law, if the Court overlooked controlling decisions or factual matters before it, if new evidence has been discovered that warrants

revisiting the prior decision, or to correct a clear error or prevent manifest injustice. *See* Local Rule 6.3; *Arthur Glick Truck Sales, Inc. v. Stephen East Corp.*, 965 F. Supp. 2d 402, 405 (S.D.N.Y. 2013); *Brown v. Barnes & Noble, Inc.*, No. 16-CV-7333 (RA) (KHP), 2020 WL 1082464, *1 (S.D.N.Y. Mar. 5, 2020). Plaintiffs argue that initial discovery has revealed that there are many internal Noom documents that contain hyperlinks to other internal Noom documents and that, even though Noom is searching for and producing relevant documents, the hyperlinked internal documents should be treated as attachments. They argue that neither the Court nor the parties appreciated that there would be thousands of Noom documents containing hyperlinks to other internal Noom documents at the time of the Court's initial ruling on this issue.

When managing discovery, the Court is guided by proportionality concerns set forth in Rule 26(b)(1) and Rule 1's mandate to ensure the just, speedy, and inexpensive determination of this action. *Fed. R. Civ. P. 1*, 26. The Court is also mindful that the parties may make early ESI decisions based on limited information, that proportionality considerations may need to be re-balanced at later points in the litigation, and that discovery plans may be modified when new information is learned. Another principle relevant to the issue before the Court is that Rule 34 requires a party to produce documents in a “reasonably usable form.” *Fed. R. Civ. P. 34*; 2006 Committee Note to *Rule 34*. This means that ESI should not be produced in a format that makes it unreasonably difficult or burdensome for the requesting party to use the documents efficiently in the litigation. *See also* Sedona Principle 12. At the same time, it is appropriate to limit collection and review to non-duplicative, relevant information. *See* Sedona Principle 8.

*4 With the above principles in mind, the Court denies Plaintiffs' motion for reconsideration. To start, the Court does not agree that a hyperlinked document is an attachment. While the Court appreciates that hyperlinked internal documents could be akin to attachments, this is not necessarily so. When a person creates a document or email with attachments, the person is providing the attachment as a necessary part of the communication. When a person creates a document or email with a hyperlink, the hyperlinked document/information may or may not be necessary to the communication. For example, a legal memorandum might have hyperlinks to cases cited therein. The Court does not consider the hyperlinked cases to be attachments. A document also may contain a hyperlink to another portion of the same document. That also is not

an attachment. A document might have a hyperlink shortcut to a SharePoint folder. The whole folder would not be an attachment. These are just examples. An email might have hyperlinks to a phone number, a tracking site for tracking a mailing/shipment, a facebook page, a terms of use document, a legal disclaimer, etc. The list goes on and on. Many of these underlying hyperlinked documents may be unimportant to the communication.

Relatedly, there has been no showing by Plaintiffs that they actually need to link to or even care about all of the hyperlinked documents. “Parties should not demand forms of production ... for which they have no practical use or that do not materially aid in the discovery process.” The Sedona Principles, Third Edition p. 173. At this point it is entirely speculative how many underlying hyperlinked documents are relevant and material to this case. While it may be true that Noom employees frequently create documents with hyperlinks to other internal Noom documents, it is not at all clear that Plaintiffs cannot identify the underlying hyperlinked documents or the quantity that are even material to this action. The process the Court already ordered is appropriate. It will allow Plaintiffs to evaluate Noom's production and, if Plaintiffs determine there is a need for an additional targeted pull or production or clarifying information about a hyperlinked document's identity or Bates number, Plaintiffs can request it.⁴ To the extent Plaintiffs have raised a concern that such a request will somehow reveal their strategy or work product, such a concern is not persuasive. It is no secret the types of documents that are relevant to this case and both sets of lawyers no doubt will have “hot docs” tags on many of the same documents in Noom's production.

⁴ The Court appreciates that there may be different versions of hyperlinked documents or authentication issues that arise. However, the number of documents where these issues may arise are likely to be a very small percentage of the overall production that can be dealt with later in discovery. The issues in this case should not raise many disputes about authenticity or versions of documents because the key issues relate to public-facing policies and procedures pertaining to Noom's autorenewal policy.

The Court has additional proportionality and [Rule 1](#) concerns: costs and delay. What Plaintiffs request is a new collection of the same emails and Google Drive documents that already

have been collected and are being reviewed for relevance and production. As noted above, Plaintiffs already agreed that Google Vault was an acceptable collection tool, at least for the Google Drive documents. Noom has argued persuasively that the redundancies of pulling hyperlinked documents would be burdensome. Indeed, one email thread may contain multiple hyperlinks to the same document that already was flagged for production. The same underlying hyperlinked document may be pulled tens if not hundreds of times in some cases. This additional collection would certainly increase the review population and, as Noom's expert explained, complicate de-duplication, delay production, and impose additional costs. Noom estimates the costs would be upwards of \$180,000. Plaintiffs have not effectively countered these projected costs. Initially, Plaintiffs posited that what they requested would cost less than \$5,000—a number that was not credibly supported by evidence. The recently submitted Declaration of Plaintiffs' ESI expert Douglas Forrest also does not adequately address the issue of costs and delay. Forrest proposes that Noom could write a program to utilize Google's Application Programming Interface (“API”) to extract links to Google Drive documents from other Google Drive documents, emails, and Slack communications. ([ECF No. 236](#).) He suggests that Noom's ediscovery programmers could create such a program within two weeks. *Id.* While it may be true that creating such a program would cost far less than \$180,000, the Forrest declaration does not address the time it would take to apply the program, load, and review the documents. Nor do Plaintiffs explain why a recollection of hyperlinked documents, many of which may be of no real value in the case and are redundant of the documents already collected, is proportional to the needs of this case.

*5 To the extent Plaintiffs argue that Noom is attempting to shield itself from discovery by virtue of its recordkeeping system, the Court rejects this argument. Noom is searching and producing its Google Drive documents—it is not refusing to search and produce them. What it objects to is collection of them through both a direct collection and a collection through hyperlinks that would dramatically increase redundancies in the collection, increase costs, and delay discovery. Moreover, Plaintiffs have not made a showing that the value of obtaining the metadata establishing the linkages for *all* hyperlinked documents is proportional to the needs of the case.⁵

⁵ Perhaps future ESI software will be able to provide greater efficiencies and reduced costs to address the concerns of both parties. Alternatively, perhaps the parties should have considered whether a

separate Google Drive collection was necessary if the documents were being collected through hyperlinks to reduce redundancies.

The Court acknowledges the inherent tension between the value and efficiency in creating an ESI protocol up front that addresses all potentially foreseeable issues on the one hand, and getting discovery underway in a case on the other hand. The existing protocol together with the Court's work around to address Plaintiffs' concerns for those key documents where they need more information about the hyperlinked documents strikes an appropriate balance based on the needs of this case. And, as reported at the most recent conference, the parties have already successfully utilized the Court's procedure.

To the extent Defendants request attorneys' fees in connection with opposing Plaintiffs' motion, their request is denied. The issue Plaintiffs raise is an important one and one on which the Court did not issue a fulsome written decision, instead opting to address the issue more informally as is the practice

in this District. Nonetheless, the Court cautions Plaintiffs not to repeatedly raise the same discovery issue in the hopes of getting a different answer absent good cause or a showing that the Court has not appreciated or misconstrued the facts or the law.

CONCLUSION

For the reasons set forth above, Plaintiffs' letter motion at ECF No. 214 is DENIED. Defendants' related request for attorneys' fees generated in responding to Plaintiffs' letter motion is also DENIED.

SO ORDERED.

All Citations

Slip Copy, 2021 WL 948646

End of Document

© 2023 Thomson Reuters. No claim to original U.S. Government Works.

EXHIBIT 2

Defendants' ~~Plaintiffs~~ Defendants' Draft – 34/4/~~32830~~/23

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF TEXAS
HOUSTON DIVISION**

IN RE APACHE CORP. SECURITIES
LITIGATION

Case No. 4:21-cv-00575

District Judge George C. Hanks, Jr.

Magistrate Judge Andrew M. Edison

CLASS ACTION

**DISCOVERY PROTOCOL REGARDING
PRODUCTION OF ESI AND PAPER DOCUMENTS**

The Parties in the above-captioned case (the “Action”), by and through their counsel, have agreed to this Discovery Protocol Regarding Electronically Stored Information and Paper Documents (the “Protocol”) to govern the searching, identification, processing, and production of hard copy (“Hard Copy Materials”) and electronically stored documents and information (“ESI,” and together with Hard Copy Materials, “Discovery Material”) in the Action.

I. GENERAL PROVISIONS

1. Nothing in this Protocol is intended to alter the Parties’ rights under or obligations to comply with any applicable Federal Rules of Civil Procedure, Local Rules of the United States District Court for the Southern District of Texas (the “Local Rules”), or the Court’s Court Procedures (the “Procedures”) regarding the production of Hard Copy Materials or ESI.

2. No provision of this Protocol waives any objections as to the relevance, discoverability, admissibility, or confidentiality of any particular Discovery Material produced in this Action. For the avoidance of doubt, a Party’s compliance with this Protocol will not be

~~Defendants' Plaintiffs~~ Defendants' Draft – 34/4/32830/23

interpreted to require disclosure of information potentially protected by the attorney-client privilege, the work product doctrine, or any other applicable privilege.

3. To the extent that any Party collected and processed documents in response to a discovery request in this ~~A~~ action before the Parties' agreement to this Protocol, and production of such documents was not made in accordance with the terms of this Protocol, the Parties will meet and confer if the other Parties have questions or concerns about the format of any such productions.

4. References to deadlines in this Protocol shall comply with Federal Rule of Civil Procedure 6 with respect to computing deadlines.

5. Production of responsive, non-privileged documents shall be completed on a rolling basis. The Parties agree that Documents may be produced without a prior meet-and-confer.

6. The Parties may mutually modify or revise this Protocol by agreement as circumstances dictate ~~or as ordered by the Court.~~

7. The Parties shall meet and confer in an effort to resolve any disputes that may arise under this Protocol before seeking relief from the Court.

II. DEFINITIONS

1. "Archival systems" mean long-term repositories for the storage of records that preserve the content of the documents or data stored therein, prevent or track alterations to such documents or data, and control access to electronic records.

2. "Back-up systems" refers to computer systems used to store copies of electronic information on back-up media (magnetic tapes, CDs, DVDs, hard drives, or other storage), to permit recovery of the information in the event of loss or damage to the original data.

3. "Custodial Data Sources" means any potentially relevant data sources unique to any particular individual that are ~~either~~ in the possession, custody, or control of ~~or reasonably accessible to a Party and reasonably likely to contain responsive information.~~

Commented [A1]: As Defendants have made clear, we are willing to agree to a voluntary ESI protocol agreement among the parties, but we do not agree to entry of a court order on these matters.

~~Defendants' Plaintiffs~~ Defendants' Draft – 34/4/32830/23

but are not limited to, any applicable local hard drives, network home or personal file shares, removable storage, email, online or cloud storage (e.g., OneDrive, SharePoint), removable storage media, messaging applications or software, communication, or coordination platforms (e.g., Microsoft Teams), phones, tablets, laptops, and physical files.

4. “Document” shall have the meaning contemplated by Federal Rule of Civil Procedure 34(a)(1)(A), which includes Electronically Stored Information (“ESI”).

5. “Electronic message” means any electronic text or media content exchanged between two or more users of a software application, including but not limited to SMS messages sent over cellular networks and messages sent over the Internet using WhatsApp, iMessage, and Microsoft Teams (instant messaging), among other applications.

6. “Email” means an electronic means for communicating written information that is deliverable to designated recipients at specific addresses associated with particular Internet domains.

7. “Extracted text” means the text extracted from a Native document, and includes all header, footer and document body information when available.

8. “Load file” means an electronic file used to import required production information into a document database, including, document images, extracted text or OCR text, native files, and metadata, as well as information indicating document breaks and relationships (e.g., parent-child relationships for email families).

9. “Metadata” means structured data about ESI that is created by the file system or application, embedded in the document, and as modified through ordinary business use, that describes the characteristics, origins, usage, and validity of the ESI, among other things.

10. “Native format” means the format of ESI in the application in which such ESI was

~~Defendants'~~ ~~Plaintiffs~~ ~~Defendants'~~ Draft – 34/4/32830/23

originally created.

11. “Network” or “shared storage systems” shall mean any data storage device accessible to multiple users remotely over a computer network.

12. “Non-Custodial Data Sources” means any potentially relevant enterprise, network or shared data sources accessible by multiple individuals that are in the possession, custody, or control of a Party and may include, but are not limited to, any applicable databases, file servers, storage area networks (SANs), network-attached storage (NASs), email servers or platforms, web servers or platforms, communication or coordination platforms (e.g., Microsoft Teams), third-party-hosted Software as a Service (SaaS) platforms, document management systems (DMS), record management systems (RMS), content management systems (CMS), departmental/project/collaborative/shared storage spaces, e-rooms, structured data stores, application data, online or cloud storage (e.g., OneDrive, SharePoint), ~~back-up systems or media,~~ ~~archival systems,~~ and hard-copy document repositories.

13. “OCR” means the optical character recognition technology used to read hard copy paper documents or electronic images of documents, and output such documents to a searchable text format.

14. “Producing Party” means any Party that produces documents or ESI pursuant to this Protocol.

15. “Receiving Party” means any Party that receives documents or ESI produced in this Action.

16. “Requesting Party” means any Party that requests production of documents or ESI in this Action.

17. “Responsive document” means any document or ESI that a Party produces in

~~Defendants' Plaintiffs~~ Defendants' Draft – 34/4/32830/23

response to any discovery request or subpoena served in this Action, subject to any limitations under the Federal Rules of Civil Procedure, Local Rules, Procedures, or order of this Court.

18. “Search term” means a word or a combination of words or phrases designed to capture potentially relevant ESI and includes strings of words and phrases joined by proximity and Boolean connectors or other syntax.

19. “Structured data” means data that resides in a fixed field within a record or file, or stored in a structured format, such as databases (such as Oracle, SQL, Access) or data sets, according to specific form and content rules as defined by each field of the database.

20. “Tagged Image File Format” or “TIFF” refers to the Group IV graphic file format for storing bit-mapped images of documents and ESI.

21. “TAR” (Technology-Assisted Review) means a process for prioritizing or coding a collection of ESI using a computerized system, using algorithms or systematic rules, that harnesses human judgments of subject matter expert(s) on a smaller set of documents and then extrapolates those judgments to the remaining documents in the collection. TAR systems generally incorporate statistical models and/or sampling techniques to guide the process and to measure overall system effectiveness.

22. “Text file” is a file containing the full text of native files extracted directly from the native file, or, in the case of hard copy documents or scanned PDF documents, subject to OCR, a file containing the text resulting from the OCR.

23. “Unstructured data” refers to free-form data which either does not have a data structure or has a data structure not easily readable by a computer without the use of a specific program designed to interpret the data, including but not limited to, word processing documents, slide presentations, email, PDFs, spreadsheets, and webpages, blogs, image files, instant messages,

Defendants' ~~Plaintiffs~~ Defendants' Draft – 34/4/32830/23

audio and video files, and others of similar variable format.

III. COLLECTION AND PROCESSING

1. The parties shall use methods of collection and processing (a) that preserve the integrity of document metadata, (b) that preserve parent-child and family group relationships such as the association between parent documents and attachments, ~~(c) that preserve the relationship between emails, electronic messages, chats, or channel messages (e.g., within Microsoft Teams) that contain links to cloud attachments (e.g., within SharePoint or Teams) and the cloud attachments so linked,~~ and (c) ~~de~~ that extract embedded files from Documents, subject to paragraph III.C below.

2. **Discovery Not Required.** Except as otherwise agreed to by the Parties or ordered by the Court, ~~T~~he Parties need not collect the following ESI:

- a. Data stored on photocopiers, scanners, and fax machines.
- b. Any electronic file which matches the Reference Data Set published by the National Institute of Standards and Technology's National Software Reference Library.
- c. System or executable files (.exe, .dll, etc.).
- d. Structural files not material to individual file contents that do not contain substantive content (.css, .xsl, .xml, .dtd, etc.).
- e. Documents from data sources that a party claims are inaccessible due to undue burden or cost.
- f. Backup data files that are maintained in the normal course of business for purposes of disaster recovery, including backup tapes, discs, SAN, and other forms of media, and are substantially duplicative of data that is more accessible elsewhere.

Commented [A2]: A gain, what you are describing in your proposed additions are NOT attachments. Therefore, this language does NOT belong in here. To the extent there is data available in its current form and captured by our collection, we will obviously review it. We will also be reviewing SharePoint and OneDrive data for potential production. We cannot, however, agree to preserve attachments that do not exist, which is what your language contemplates.

~~Defendants' Plaintiffs~~ Defendants' Draft – 34/4/32830/23

- g. Deleted, slack, fragmented, or unallocated data.
- h. Random access memory (RAM), temporary files, or other ephemeral data that are difficult to preserve without disabling the operating system.
- i. On-line access data such as temporary internet files, history files, cache files, and cookies.
- j. Data in metadata fields frequently updated automatically, such as last-opened or last-printed dates, except to the extent specified in the Metadata Field Table attached hereto.
- k. Data remaining from systems no longer in use that are unreadable on the systems in use.
- l. Other standard system file extensions, including but not limited to BIN, CAB, CHK, CLASS, COD, COM, DLL, DRV, INF, INI, JAVA, LIB, and SYS.

B. Embedded Objects. Logos, icons, and footers that are embedded in Documents may be culled from a data set and need not be produced as separate documents by a Producing Party (e.g., such embedded objects will be produced within the document itself, rather than as separate attachments).

IV. SEARCH METHODOLOGY

A. Transparency and Cooperation. The Parties agree to be reasonably transparent regarding the universe of documents subject to collection and the methodologies being used to identify responsive documents. If, after reviewing Discovery Materials received from the Producing Party, the Requesting Party believes that additional searches or collections may be necessary, it shall request a meet-and-confer to determine what, if any, additional searches the

Defendants' ~~Plaintiffs~~ Defendants' Draft – 34/4/32830/23

Producing Party is willing to conduct.

B. Custodians. Each Producing Party will disclose a list of the document custodians whose data will be searched for responsive documents, as well as the Custodial Data Sources being searched for each custodian.

C. Non-Custodial Sources. Each Producing Party will disclose a list of the Non-Custodial Document Sources being searched.

D. Search Terms. Where a Producing Party elects to use search terms to identify documents and ESI potentially responsive to a Party's requests for production, the Producing Party will provide a list of proposed search terms that it believes would lead to the identification of responsive documents and ESI. Within 5 business days of receiving the Producing Party's proposed search terms, the Requesting Party will provide any proposed revisions to the terms. Within 5 business days of receipt of the Requesting Party's proposed revisions, the Producing Party will identify the terms that it is willing to run and any modifications that it is willing to make to its proposed search terms in light of the Requesting Party's proposed revisions. If disputes still exist at the end of that process, the Parties may submit them to the Court pursuant to the Court's practices and procedures for raising discovery disputes.

E. Technology-Assisted Review. Where a Producing Party elects to use predictive coding/technology-assisted review, ~~instead of search terms,~~ to identify documents and ESI potentially responsive to a Party's requests for production, the Producing Party will disclose (i) the data sources against which TAR will be run; (ii) the TAR or advanced analytics methodology deployed; (iii) the quality control measures used to validate the results, including recall and precision rates; (iv) any documents, document types, file types, or other categories of ESI to be excluded from the TAR process; and (v) the method used to identify such documents for exclusion.

Commented [A3]: Neither the rules nor the Sedona Principles permit one party to dictate how another performs its review. We will determine the appropriate process based on volume of data to be reviewed, we will be transparent about the process if we elect to go to TAR, and there is a process in here for plaintiffs to raise concerns.

~~Defendants' Plaintiffs~~ Defendants' Draft – 34/4/32830/23

If any Party has concerns about another Party's TAR protocol, the Parties will meet and confer to try to address them. If disputes still exist at the end of the meet and confer process, the Parties may submit them to the Court pursuant to the Court's practices and procedures for raising discovery disputes. ~~The parties agree that a producing party may not to utilize TAR after application of search terms except to prioritize the review of all documents hitting on the agreed upon search terms.~~

~~E. — Validation. The Parties will meet and confer regarding appropriate measures for validating the results of the methodologies applied to identify potentially responsive ESI.~~

V. PRODUCTION OF HARD COPY DOCUMENTS

Hard copy documents should be scanned as single-page, Group IV, 300 DPI TIFF images and produced with an .opt image cross-reference file and a delimited database load file (*i.e.*, .dat). The database load file should contain the following fields: "BEGBATES," "ENDBATES," "PAGES," "VOLUME," and "CUSTODIAN." The documents should be logically unitized (*i.e.*, distinct documents shall not be merged into a single record, and single documents shall not be split into multiple records) and be produced in the order in which they are kept in the usual course of business. If an original document contains color, the document shall be produced as single-page, 300 DPI JPG images with JPG compression and a high-quality setting as to not degrade the original image. Multi-page OCR text for each document should also be provided. The OCR software shall maximize text quality. Settings such as "auto-skewing" and "auto-rotation" should be turned on during the OCR process. Each text file shall be named with the beginning Bates number of the hard copy document to which the text file relates. The parties agree that should the volume of the Hard Copy Materials make the production of single page 300 dpi JPG images with JPG compression unreasonably burdensome, the parties will meet and confer in efforts to reduce the volume of documents to be produced as single page 300 dpi JPG images with JPG compression (as opposed to black and white TIFF images).

Defendants' ~~Plaintiffs~~ Defendants' Draft – 34/4/32830/23

VI. PRODUCTION OF ESI

A. **Format.** All documents and ESI produced in this Action shall be produced in the formats described below:

a. **Images.** Except as otherwise specified herein, all other documents and ESI shall be produced as single-page, black and white, TIFF Group IV, 300 DPI TIFF images. If an original document contains color and is an image-based file (e.g., .jpeg), or a PowerPoint, or color is material to the understanding of the document, then the document should be produced as single-page, 300 DPI JPG images with JPG compression and a high-quality setting as to not degrade the original image. TIFFs/JPGs shall be processed to show any and all text and images that would be visible to the reader using the native software that created the document (e.g., TIFFs of emails should include the BCC line). To the extent that Excel files and other spreadsheets and other files that are not able to be imaged must be imaged to redact privileged information under the terms herein, they shall be processed in a manner that maintains and displays all hidden columns or rows, hidden text or worksheets, speaker notes, track changes, and comments. However, the Parties are under no obligation to enhance an image beyond how it was kept in the usual course of business. If the image does not accurately reflect the document as it was kept in the usual course of business, the Parties agree to meet and confer in good faith on production format options. Page orientation for TIFF/JPG images shall be the same as the underlying document from which the image is created.

b. **Native Files.** Microsoft Excel and other spreadsheet file types, including comma or tab delimited text files, PowerPoint, and other presentation file types, Microsoft Access files, and audio and video files, photo or graphic images, documents with tracked changes in the metadata, and other file types that – when produced in a non-native format

Defendants' ~~Plaintiffs~~ Defendants' Draft – 34/4/32830/23

– limit or lose relevant content, shall be produced in native format. For each document or ESI required to be produced in native format, the Producing Party shall also provide a single-page Bates stamped image slip sheet stating the document has been “PRODUCED IN NATIVE FORMAT.” Each native file should be named according to the Bates number it has been assigned and should be linked directly to its corresponding record in the load file using the NATIVELINK metadata field. PowerPoint documents shall be produced in native format along with single-page, 300 DPI TIFF/JPG color-for-color images, as applicable. In the event that any document or ESI produced in image format is: (i) unreadable or illegible; (ii) cuts off or is missing data or information; or (iii) cannot be utilized at a deposition, a Receiving Party may request that such file be produced in native format in accordance with the procedures outlined in this paragraph, and the Producing Party will comply with all such reasonable requests. To the extent a Receiving Party believes that any other documents or ESI should be produced in native format, it will identify such documents or ESI and the basis for the request, and the Producing Party will either produce the native files or explain why it refuses to produce the native files, and the Parties agree to meet and confer regarding such requests.

c. **Electronic messages.** Electronic messages (e.g., iMessage, SMS, MMS, Teams Chat) shall be produced in a searchable format that preserves the presentational features of the original message, such as emojis, images, video files, animations, and the like to the extent reasonably possible, and shall be produced with its relevant conversation thread broken down by conversation day, along with relevant contextual messages a Producing Party reasonably identifies. To the extent a Requesting Party reasonably believes additional contextual messages are necessary to understand a particular responsive

Defendants' ~~Plaintiffs~~ Defendants' Draft – 34/4/32830/23

message, the Requesting Party shall identify the responsive message by Bates number, and Parties will meet and confer regarding what, if any, relevant, non-privileged contextual messages exist for possible production. To the extent a Requesting Party believes that an electronic message was not produced in a searchable format that preserved the presentation features of the original message, the Requesting Party shall identify the message by Bates number, and the Producing Party will reproduce it in a searchable format that preserves the presentation features of the original message if possible.

d. **Structured Data.** To the extent a response to discovery requires production of discoverable ESI stored in a structured database, the Parties shall meet and confer in an attempt to agree upon a set of queries to be made for discoverable information and generate a report in a reasonably usable and exportable electronic file for review by the Requesting Party. Upon review of the report, the Requesting Party may make reasonable requests for additional information, including different reports or report formats, and specific data from identified fields.

e. **Proprietary Software.** To the extent that relevant ESI cannot be rendered or reviewed in a reasonably useable form without the use of proprietary software, the Parties will meet and confer to determine whether and to what extent ~~identify an effective solution to~~ producing such ~~relevant~~ ESI is necessary and, if so, how that can be accomplished. ~~Any costs associated with production of ESI subject to this provision shall be borne by the Requesting Party.~~

f. **Attachments.** If any part of an email or its attachments is responsive, the entire email and all attachments shall be produced, except: (i) where an attachment is withheld pursuant to a valid claim of privilege, in which case that attachment will be

~~Defendants'~~ ~~Plaintiffs~~ ~~Defendants'~~ Draft – 34/4/32830/23

replaced with a single-page TIFF image identifying the privilege or work product protection being claimed, and the rest of the family produced; and (ii) where the attachment is a non-substantive, automatically generated embedded file (e.g., logos) or formatting file (e.g., .ole or .dll). Attachments shall be produced sequentially after the parent email. For such attachments, the BEGATTACH and ENDATTACH metadata fields shall be populated to show the family relationship.

~~Modern or Cloud Attachments/Hyperlinked Documents. Documents sent via links to internal or non-public documents (e.g., to Microsoft 365, SharePoint, etc.) that are contained within an email, message, chat or other document, shall be extracted and produced as an attachment, with the parent-child relationship preserved and Metadata fields showing the family relationship and FilePath to such documents populated.~~

g. **Embedded Files.** Except as set forth above in § III.C, embedded documents or files embedded in documents (e.g., embedded MS Office files) and images embedded in RTF files shall be extracted as separate files and treated as attachments in accordance with the procedures described herein in the prior subsection (§VII.A.f) with the parent/child relationship preserved.

h. **Dynamic Fields.** All dynamic date and time fields, where such fields are processed to contain a value, shall, to the extent reasonably possible, be processed with a single date and time setting that is consistent across each Party's productions, Coordinated Universal Time (UTC), and in a way that maintains the date/time shown in a document as it was last saved by the custodian or end user, not the date of collection or processing (i.e., force off auto data).

i. **Hidden Content, Speaker Notes, Comments, and Tracked Changes.** To

~~Defendants' Plaintiffs~~ Defendants' Draft – 34/4/32830/23

the extent that a document contains hidden content, tracked changes, comments, notes, or other similar information, it shall be imaged so that such information is captured on the produced image file as it would appear in the native format of the document subject to any claim of privilege or work product protection, in which case redactions will be applied.

j. **Compressed File Types.** Compressed file types (*e.g.*, .ZIP, .RAR, .CAB, .GZ) should be decompressed so that the lowest level document or file is extracted.

k. **Encrypted Files.** Certain files may contain encrypted data or be password protected. Encrypted data and password protected files will be produced only where the key or password can be ascertained after reasonable efforts. To the extent an encrypted or password protected file is inaccessible to the Producing Party but deemed potentially responsive, to the extent available and not subject to a claim of privilege, the Party will identify the file attributes (directory, size, name, extension, last modification date) of any such inaccessible encrypted or password protected files.

B. De-Duplication. Prior to production, each Party will remove exact duplicate stand-alone documents and document families, as identified based on ESI using hash value matching (such as MD5 or SHA-1 hash values). Documents and document families that are not duplicates may not be removed. No Party shall identify and/or eliminate electronic duplicates by manual review or some method other than by use of the technical comparison using MD5 or SHA-1 hash values outlined above. The Parties will de-duplicate globally across the document collection. Duplicate documents held in different file path locations and/or by different custodians do not need to be produced, provided that the load file produced must include a Custodian-All metadata field that shows each custodian from whom or which a copy of the produced document was collected. The Parties shall de-duplicate stand-alone documents against stand-alone documents and shall de-

~~Defendants' Plaintiffs~~ Defendants' Draft – 34/4/32830/23

duplicate entire document families against entire document families, in a manner that does not break apart families and shall be performed at a family level.

C. Email Threading. Where an email chain consists of more than one segment, the Producing Party may use industry-standard email threading technology such that only the most inclusive message(s) (i.e., those containing a full and complete reflection of all the text and attachments contained anywhere within the email chain), any lower-included emails that have attachments, and any lower-included email that contains unique content not included elsewhere in a produced email string must be produced. ~~To the extent a Producing Party applies email threading, it shall produce metadata (including email header information) for all lesser included emails that are not independently produced.~~ To the extent a Receiving Party reasonably believes that a lesser-inclusive email should be produced, it may identify those documents by Bates number, and the Parties agree to meet and confer regarding production.

D. Redactions. A Party may only apply redactions to protect attorney-client privilege; attorney work product; joint defense privilege; highly confidential business information (as described below); information prohibited from disclosure by federal, state, or foreign statutes or regulations; personal health information concerning any individual and personal identity information that a Party is required not to disclose by federal, state or foreign statute, or regulation. Redactions shall be limited to the specific privileged information protected and shall not include email header information (unless itself privileged), or factual or contextual information contained in the document. To address particularized concerns regarding the production of sensitive business information contained in the Company's Board materials, the parties agree that a Party also may redact specific highly confidential business information relating to non-relevant businesses contained in Board materials only if (1) the redacted information is both nonresponsive and

~~Defendants'~~ ~~Plaintiffs~~ ~~Defendants'~~ Draft – 34/4/32830/23

irrelevant; and (2) the redacted information would not assist in understanding or providing context for the relevant portion of the document or document family of which it is a part. Such redactions must be identified as “Redacted – Highly Confidential Nonresponsive Information.” To the extent a Party is required to withhold from production a member (email or attachment) of an email family on the basis of privilege including work production protection, where some or all of the other email family members are being produced, the producing Party shall provide a Bates numbered placeholder TIFF image bearing the legend, “DOCUMENT WITHHELD FOR PRIVILEGE.” To the extent a Party is required to Redact a document required to be produced in Native format under this agreement, it shall to the extent reasonable redact the document using a native redaction tool and, where native redaction is not reasonable, produce the document in a manner that displays all non-privileged information that would be available in the document’s native format, including hidden rows and columns for excels and other spreadsheets, speaker notes and comments for Power Points and other presentations, and track changes and comments for Word documents, as well as the date, time and author of any such changes and comments. Redactions will be clearly indicated. For redacted documents, a text file shall be created using OCR and shall be produced in lieu of extracted text. The searchable text file will include the word “Redacted.”

E. Bates numbering. Each TIFF/JPG image produced shall be branded with a Bates number that (i) identifies the Producing Party; (ii) maintains a constant length of nine numeric digits (including 0-padding) across the entire production; and (iii) is sequential, both within a given document, but should not include the Bates number in the extracted text of the ESI. Each TIFF/JPG image file shall be named with the same page-level Bates number branded on the underlying image. For ESI that is produced subject to a claim of confidentiality pursuant to the Protective Order entered in this Action (the “Protective Order”), the Producing Party shall electronically

Defendants' ~~Plaintiffs~~ Defendants' Draft – 34/4/32830/23

brand the appropriate confidentiality designation onto each page of the document, but should not include the confidentiality designation in the extracted text of the ESI.

F. Text Files and OCR. All documents and ESI shall be produced with a corresponding multipage searchable text file (i.e., one .TXT file per electronic file as opposed to one text file per page). The text file for ESI shall contain full text extraction and be created by extracting text directly from the underlying native file, unless the ESI must be redacted prior to production, in which case the text file shall be generated by applying industry standard OCR technology to the redacted version of the ESI. Each text file shall be named with the beginning Bates number of the electronic file to which the text file relates.

G. Load files. All documents should be provided with Concordance-compatible image and data load files (i.e., .OPT and .DAT files) using standard Concordance delimiters showing the Bates number of each page and the approximate unitization of the documents.

H. Metadata. All ESI will be produced with a delimited, database load file that contains the metadata fields listed in Table 1, attached hereto. The metadata produced should have the correct encoding to enable preservation of the documents' original language. For ESI other than email and e-docs that do not conform to the metadata listed in Table 1, such as electronic messages, and ESI from collaboration tools, etc., the Parties will meet and confer as to the appropriate metadata fields to be produced.

I. Production Media. All documents and ESI shall be produced via secure FTP or other secure file transfer or other mutually agreeable production media, along with an identification of the Bates range of the production. Any replacement productions shall be clearly identified as such. All productions shall be encrypted, and the Producing Party shall provide a decryption key at the time of production, via separate cover.

Defendants' ~~Plaintiffs~~ Defendants' Draft – 34/4/32830/23

J. Privilege Log. The Parties agree to exchange privilege logs that comply with Federal Rule of Civil Procedure 26(b)(5) on a rolling basis, with the final privilege log being produced no later than thirty (30) days after the Producing Party has substantially completed its document production. Privilege logs will be provided in a searchable and sortable spreadsheet format. The following privileged communications or work product need not be included in a privilege log unless otherwise agreed to by the Parties or ordered by the Court: (a) communications to or from outside legal counsel related to this Action; (b) work product of counsel and Parties related to this Action; (c) any privileged communications that post-date the filing of the initial complaint in this Action; and (d) any communications regarding litigation holds or preservation, collection, or review of discoverable information in this Action. With respect to e-mail threads for which a Party claims a privilege, over one or more of the discrete communications within the thread, only the most inclusive e-mail ~~shall~~ need be logged, but the log will set forth the basis for the privilege claim for each lesser inclusive email within the thread for which a privilege claim is made regardless of the segment in which the privileged information exists within the inclusive e-mail, and shall provide metadata for each of the lesser inclusive emails for which a claim of privilege is made. The Parties will meet and confer in good faith regarding the scope of the privilege log should any such issues arise. Further, the Parties agree that documents that have been redacted for privilege should also be logged separately and include the fields as set forth in the following paragraph.

In an effort to avoid unnecessary expense and burden, the Parties agree that, for documents withheld from production or redacted on the basis of attorney-client privilege, work product doctrine, and/or any other applicable privilege, the Producing Party will prepare a log containing, for each document (except those exempted above): (a) the privilege(s) claimed; (b) the date of the

Defendants' ~~Plaintiffs~~ Defendants' Draft – 34/4/32830/23

document; (c) the author(s), addressee(s), and all recipients of the document and, to the extent available in the metadata for the logged document, the email addresses for all such persons; (d) a description of the general subject matter of the privileged information contained in the document sufficient to allow the Receiving Party (and the Court as necessary) to assess the privilege claimed; and (e) the custodian. The log will be accompanied by a list identifying any person on the log who is an attorney or an employee of any legal department. For redacted documents, if there is more than one redaction per document, one log entry can be used to describe the multiple redactions.

If the Receiving Party reasonably requires further information, it shall explain in writing the need for such information and identify, by Bates number or other unique identifier, each document for which it seeks this information. The Producing Party will respond to any such reasonable request by providing a more detailed description of the nature of the withheld information that will enable the Receiving Party to assess the claim of privilege.

If the Receiving Party contests a privilege claim, the Receiving Party must notify the Producing Party in writing. The Parties must attempt to confer in good faith to resolve any challenge. Thereafter, the Parties may seek judicial intervention if the challenge cannot be resolved.

K. Privilege Claw Back. The inadvertent production of privileged or work-product protected data (including metadata) is governed by the applicable Federal Rules of Civil Procedure and the Protective Order entered in the Action.

Dated: _____

Counsel for Plaintiffs

Dated: _____

Counsel for Defendants

Defendants' ~~Plaintiffs~~ Defendants' Draft – 34/4/32830/23

TABLE 1
Metadata Fields¹

Field Name	Example/Format	Description
BEGBATES	ABC0000001 (Unique ID)	The Document ID number associated with the first page of a document.
ENDBATES	ABC0000003 (Unique ID)	The Document ID number associated with the last page of a document.
BEGATTACH	ABC0000001 (Unique ID Parent-Child Relationships)	The Document ID number associated with the first page of the parent document.
ENDATTACH	ABC0000008 (Unique ID Parent-Child Relationships)	The Document ID number associated with the last page of the last attachment.
VOLUME	VOL001	The name of CD, DVD, or Hard Drive.
RECORDTYPE	Options: eMail, Attachment, Scanned Doc, eFile	The record type of a document.
SENTDATE	MM/DD/YYYY	The date the email or calendar entry was sent.
SENTTIME	HH:MM	The time the email or calendar entry was sent.
RECEIVEDDATE	MM/DD/YYYY	The date the document was received.
RECEIVEDTIME	HH:MM	The time the document was received.
CREATEDATE	MM/DD/YYYY	The date the document was created.
CREATETIME	HH:MM	The time the document was created.
LASTMODDATE	MM/DD/YYYY	The date the document was last modified.
LASTMODTIME	HH:MM	The time the document was last modified.
MEETING START DATE	MM/DD/YYYY	Start date of calendar entry.
MEETING START TIME	HH:MM	Start time of calendar entry.
MEETING END DATE	MM/DD/YYYY	End date of calendar entry.
MEETING END TIME	HH:MM	End time of calendar entry.
<u>FILEPATH</u>	<u>i.e. /JsmithPC/Users/Jsmith/Desktop</u>	The file path from the location in which the document was stored in the usual course of business. This field should be populated for both email and e-files.
AUTHOR	jsmith	The author of a document from extracted metadata.
LASTEDITEDBY	jsmith	The name of the last person to edit the document from extracted metadata.
FROM	Joe Smith <jsmith@email.com>	The display name and email address of the author of an email/calendar item. An email address should always be provided.
TO	Joe Smith <jsmith@email.com>; tjones@email.com	The display name and email address of the recipient(s) of an email/calendar item. An email address should always be provided for every email if a recipient existed.

Defendants' ~~Plaintiffs~~ Defendants' Draft – 34/4/32830/23

CC	Joe Smith <jsmith@email.com>; tjones@email.com	The display name and email of the copyee(s) of an email/calendar item. An email address should always be provided for every email if a copyee existed.
BCC	Joe Smith <jsmith@email.com>; tjones@email.com	The display name and email of the blind copyee(s) of an email or calendar item. An email address should always be provided for every email if a blind copyee existed.
SUBJECT		The subject line of the email/calendar item.
MESSAGE TYPE	Appointment, Contact, Task, Distribution List, Message, etc.	An indication of the email system message type.
IMPORTANCE	Normal, Low, High	Email Importance Flag
TITLE		The extracted document title of a document.
CUSTODIAN		The custodian of a document from which the document originated.
CUSTODIAN-ALL	Smith, Joe; Doe, Jane	All of the custodians of a document from which the document originated, separated by semicolons.
ATTACH COUNT	Numeric	The number of attachments to a document.
FILEEXT	XLS	The file extension of a document.
FILENAME	Document Name.xls	The file name of a document.
FILESIZE	Numeric	The file size of a document (including embedded attachments).
IS EMBEDDED	Yes or No	The yes/no indicator of whether a file is embedded in another document.
HASH		The MD5 or SHA-1 Hash value or "de-duplication key" assigned to a document. The same hash method (MD5 or SHA-1) should be used throughout production.
REDACTED	Yes or No	The yes/no indicator of whether a file contains redactions.
CONVERSATION INDEX		ID used to tie together email threads.
TIMEZONEPROCESSED	PST, CST, EST, etc	The time zone the document was processed in. NOTE: This should be the time zone where the documents were located at time of collection.
NATIVELINK	D:\NATIVES\ABC000001.xls	The full path to a native copy of a document.
FULLTEXT	D:\TEXT\ABC000001.txt	The path to the full extracted text of the document. There should be a folder on the deliverable containing a separate text file per document These text files

¹For any ESI that does not conform to the metadata listed here, the parties will meet and confer as to the appropriate metadata fields to be produced.